Review of Probability

Models ch. 1-2
Simulation ch. 2 - nice, concise review

Outline

1 - Sample spaces & events
   - Conditional probability
   - Bayes' Formula
   - Independence

2 - Random Variables
   - Discrete/Continuous RVs
   - Common Distributions (Table on website)
   - Expectation, Variance
   - Joint Distributions
   - Moment generating functions
   - LLN, CLT, other limit theorems
Sample Space & Events

Consider an experiment (outcome not known in advance).

Let \( S \) be the \underline{sample space} of the experiment:
the set of all possible outcomes.

\[ \text{Examples:} \]

- Toss a coin: \( S = \{ H, T \} \)
- Toss a coin twice: \( S = \{ HH, HT, TH, TT \} \)
  \hspace{1cm} \text{ordered, here all outcomes are equally likely (all outcomes have prob. \( \frac{1}{4} \))}
- Roll a die:
  \[ S = \{ 1, 2, 3, 4, 5, 6 \} \]
- Roll a pair of dice:
  \[ S = \{ (i, j) : i, j \in \{1, \ldots, 6\} \} \]
  \[ |S| = 36 \text{ equally likely outcomes} \]

\[ \text{def: Event} = \text{subset of } S \]

\[ \text{e.g. event } A = \{ 2, 4, 6 \} \text{ roll an even \#} \]

Operations on events (sets)

\[ \text{union } \cup \]
\[ \text{intersection } \cap \]
\[ \text{complement} \]
\[ \text{mutually exclusive events (aka disjoint): } A \cap B = \emptyset \]
Axioms of Probability

For each event $A$ of $S$, the probability of $A$ denoted $P(A)$ satisfies the following axioms:

1) $0 \leq P(A) \leq 1$

2) $P(S) = 1$

3) For any sequence of mutually exclusive events $A_1, A_2, \ldots$ (i.e. $A_i \cap A_j = \emptyset$ for $i \neq j$)

$$P\left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} P(A_i) \quad \text{for } n=1,2,\ldots$$

Note: $1 = P(S) = P(A \cup A^c) = P(A) + P(A^c)$

always mut. exclusive

$$\Rightarrow P(A^c) = 1 - P(A) \quad \text{very useful property!}$$

So $P(\emptyset) = 1 - P(S) = 0$

Note: Events $A \cap B$ on $S$.

$A \cap B \Rightarrow P(A) \leq P(B)$

Addition Rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Conditional Probability

Events \( A \cap B \) on sample space \( S \).

\[
P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{for } P(B) > 0
\]

\[
\Rightarrow P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)
\]

"Multiplication Rule"

Intuition:

\[ S \]

- Universe (all possible outcomes) = \( S \)
- \( P(A) = \# \text{ of outcomes in } A \)
- \( \# \text{ outcomes in } S \)

\[ S \]

- Same idea for \( P(B) \)

Now suppose we know event \( B \) has occurred.
What's the probability \( A \) will occur? \( P(A \mid B) \)

\[ S \]

Now the universe is \( B \).
Looking for \( P(A \cap B) \) in that
univ. So we divide by \( P(B) \)

\[
:\Rightarrow \ P(A \mid B) = \frac{P(A \cap B)}{P(B)}
\]
Example: Roll 2 dice. Observe that the 1st one is 5.
What is the probability that the sum of the two dice is 7?

B: 1st die is 5
Find P(A | B).

\[ A: \text{sum is 7} \]

Possible outcomes: \((5,1), (5,2), (5,3), (5,4), (5,5), (5,6)\)

only one that sums to 7

so \(P(A | B) = \frac{1}{6}\)

\[ P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{1}{6}} = \frac{6}{36} = \frac{1}{6} \]

Q: What if \(A: \text{sum is } \leq 7\)? \(P(A \cap B) = \frac{2}{36}\)
\[ \Rightarrow P(A | B) = \frac{1}{3} \]

Independence

def: Two events \(A \& B\) are independent if

\[ P(A \cap B) = P(A) \cdot P(B) \]

Also,
\[ P(A | B) = P(A) \]

conditioning on \(B\) has no effect on prob. of \(A\)
Bayes' Formula \text{ is Law of total prob.}

\text{def: Partitioned Sample Space: Sequence of events } B_1, \ldots, B_n
\text{ s.t. } B_i \cap B_j = \emptyset \text{ for } i \neq j
\text{ and } \bigcup_{i=1}^{n} B_i = S

\text{Law of total probability}

\begin{align*}
P(A) &= \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i) \\
&= P(\bigcup_{i=1}^{n} (A \cap B_i)) \quad \text{since } (A \cap B_i) \cap (A \cap B_j) = \emptyset \text{ for } i \neq j \quad \text{must, excl. events}
\end{align*}

\text{Bayes' Formula}

\begin{align*}
P(B_j \mid A) &= \frac{P(A \cap B_j)}{P(A)} = \frac{P(A \cap B_j)}{\sum_{i=1}^{n} P(A \mid B_i) P(B_i)}
\end{align*}

\text{Inverse cond. prob. from LTP above}
Random Variables

Recall: Probability experiment (e.g. flip a coin 10 times)
large sample space $S$ (describe $S$)

often we are interested in the value of some numerical quantity determined by the outcome of exp.
(e.g. # of heads in 10 tosses)

described by a random variable

def: A random variable (RV) $X$ is a function from $S$ to $\mathbb{R}$, i.e. real-valued function.

$$X: S \rightarrow \mathbb{R}$$

eg. $X \in \{0, 1, 2, \ldots, 10\}$ for exp. above

Discrete vs. Continuous RVs

<table>
<thead>
<tr>
<th>Discrete</th>
<th>Continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>finite or countably infinite $S$</td>
<td>uncountable $S$ (i.e. interval of $\mathbb{R}$)</td>
</tr>
</tbody>
</table>
Probability density function (PDF) $p_x$ or $f_y$

Discrete RV $X$: $p_x(k) = P(X = k)$ for all $k \in S$

Continuous RV $Y$: $f_y(y)$ s.t. $P(a \leq Y \leq b) = \int_a^b f_y(y) \, dy$

Cumulative distribution function (CDF) $F_x$ or $F_y$

Discrete $X$: $F_x(k) = P(X \leq k) = \sum_{j \leq k} P(X = j)$

Continuous $Y$: $F_y(y) = P(Y \leq y) = \int_{-\infty}^y f_y(t) \, dt$ for any $y \in \mathbb{R}$

Properties

- $f(t) \geq 0 \quad \forall t$
- $\int_{-\infty}^{\infty} f(t) \, dt = 1$

CDFs

- $P(a < Y < b) = F_y(b) - F_y(a)$
- $P(Y > a) = 1 - P(Y \leq a) = 1 - F_y(a)$

Q. Relationship b/t density & dist'n functions?

density is derivative of dist'n
Common Distributions

Discrete: Bernoulli, Binomial, Geometric, Poisson, Multinomial, Neg. Binomial
Continuous: Uniform, Exponential, Gamma, Normal, Beta, Pareto

Bernoulli RV

Binary outcome: success, failure
e.g., coin flip

Let \( X = \begin{cases} 1 \text{ if success} \\ 0 \text{ if failure} \end{cases} \)

Probability mass function of \( X \):

\[
\begin{align*}
P(X = 1) &= p \quad \text{success prob} \\
P(X = 0) &= 1 - p
\end{align*}
\]

\[
\Rightarrow P(X = k) = p^k (1-p)^{1-k}, \text{ for } k = 0, 1
\]

\( X \sim \text{Bernoulli} (p) \)

Preview for later

\[
\begin{align*}
E[X] &= p \\
\text{Var}(X) &= p(1-p)
\end{align*}
\]
Binomial RV

(e.g. coin flips)

Suppose $n$ independent trials, each of which results in success w/ prob. $p$ or failure w/ prob. $1-p$.

Let $X = \#$ of successes in the $n$ trials. (e.g. # of heads in $n$ coin flips)

$X \sim \text{Binomial} \left( n, p \right)$ \quad \left( X \sim \text{Bin} \left( n, p \right) \right)

PMF: of $X$

\[
P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{for } k=0, 1, \ldots, n
\]

\[\text{Also, Binomial RV is Sum of Bernoulli RVs } \xi_i \text{i.i.d.} \]

\[
X = \sum_{i=1}^{n} \xi_i \quad X \sim \text{Bin}(n, p) \quad \xi_i \sim \text{Bern}(p)
\]

Multinomial RV

Generalizes Binomial to $r$ outcome types (like rolling an $r$-sided die $n$ times)

See formula sheet for PMF

Geometric & Negative Binomial

# trials until 1st success

# trials until $r$th success \(\rightarrow\) generalizes geometric distn

\( (\text{Also, Neg. Bin } = \sum_{i=1}^{r} \xi_i \text{i.i.d Geom}(p) ) \)

Talk about Poisson distn next week!

can be used to approximate a binomial RV when $n$ large & $p$ small

\( (X=np) \)